AD-A120 591 SECOND REPORT OF THE MULTIRATE PROCESSOR (MRP) FOR DIGITAL VOICE COMMUNICATIONS(U) NAVAL RESEARCH LAB WASHINGTON DC G S KANG ET AL. 30 SEP 82 NRL-8614 1/1

UNCLASSIFIED F/G 17/2 NL

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

NRL Report 8614

Second Report of the Multirate Processor (MRP) for Digital Voice Communications

G. S. KANG AND L. J. FRANSEN

Communications Systems Engineering Branch
Information Technology Division

AD A120591

September 30, 1982

NAVAL RESEARCH LABORATORY
Washington, D.C.

82 10 21 118

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NRL Report 8614	2. GOVT ACCESSION NO. A120 591	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) SECOND REPORT OF THE MULTIRATE PROCESSOR (MRP) FOR DIGITAL VOICE COMMUNICATIONS		5. TYPE OF REPORT & PERIOD COVERED Final report on a continuing NRL problem.
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) G. S. Kang and L. J. Fransen		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Research Laboratory Washington, DC 20375		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NRL Problem 75-0114-0-2 PE33404N Proj. X0729-CC
11. CONTROLLING OFFICE NAME AND ADDRESS		12. REPORT DATE September 30, 1982
		13. NUMBER OF PAGES 28
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Voice processing
Linear predictive coding
Prediction residual coding
Wideband voice digitizer
Embedded data

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report documents recent developments of the Navy-developed voice processor that provides compatible low-data-rate and high-data-rate voice communication: 2.4 kilobits per second (kb/s) for narrowband channels; 9.6 and 16 kb/s for wideband channels. The compatibility of the narrowband and wideband channels arises from the unique characteristic of the voice processor that the transmitted bit-stream of the 16-kb/s data contains the bit-stream of the 9.6-kb/s data as a subset. Likewise, the bit-stream of the 9.6-kb/s data also contains the bit-stream of the 2.4-kb/s data as a subset. This embedded data structure makes it possible to interconnect, without user

(Continued)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. ABSTRACT (Continued)

intervention, narrowband and wideband systems via a digital rate-converter located somewhere along the link. An advantage of this method of rate conversion is that it permits end-to-end encryption of speech data. Also during overloaded or disrupted channel conditions, communication survivability may be increased by rate reduction and/or rerouting through other available narrowband communication links. This report includes results of intelligibility and conversation tests which indicate the embedded 9.6-kb/s rate compares well with continuously variable slope delta modulation (CVSD) at 16 kb/s, and the embedded 16 kb/s performs comparably to CVSD at 32 kb/s.

CONTENTS

SECOND REPORT OF THE MULTIRATE PROCESSOR (MRP) FOR DIGITAL VOICE COMMUNICATIONS

INTRODUCTION

Since 1975 the Navy has been devising a flexible voice communication system that integrates narrowband and wideband resources into a single capability to provide satisfactory communicability over a wide range of operational conditions. In particular the communication system is designed to provide:

- Secure connectivity between wideband and narrowband users, and
- Increased system survivability for wideband users through rate reduction and rerouting.

The voice processor for this communication system, presented in this report, employs the linear predictive coder (LPC) principle to generate three data rates simultaneously: 2.4, 9.6, and 16 kilobits per second (kb/s). (In this report, a 2.4-kb/s system is referred to as a narrowband system, and a 9.6 or 16-kb/s system is regarded as a wideband system. However, some may prefer to call a 9.6 or 16-kb/s system a mediumband system.) The data rate of 2.4 kb/s is for the transmission of low quality (but highly intelligible) speech to those users who do not have access to wideband links or rely exclusively on narrowband links, such as high-frequency (HF) channels. The data rate of 9.6 or 16 kb/s is for the transmission of high-quality speech over wideband channels, such as line-of-sight radio links or well-conditioned lines.

The unique characteristic of the voice processor is that the bit-stream of the 16-kb/s data contains the bit-stream of the 9.6-kb/s data as a subset. Likewise the bit-stream of the 9.6-kb/s data also contains the bit-stream of the 2.4-kb/s data as a subset. This embedded data structure makes it possible to interconnect, without user intervention, narrowband and wideband systems via a digital rate-converter located somewhere along the link. The direct rate conversion allows end-to-end encryption of the speech bit-stream and eliminates the need of analog tandeming (and resulting speech degradation). During overloaded or disrupted channel conditions, communication survivability may be increased by rate reduction and/or rerouting through other available narrowband communication links.

The initial design of the voice processor, called the Multirate processor (MRP), was documented in Naval Research Laboratory (NRL) Report 8295 in 1979 [1]. Subsequently, the MRP algorithm operating at 2.4 and 9.6 kb/s was implemented for real-time operation on a NRL-owned microprogrammable voice processor (MVP). The MRP was extensively tested in 1980 under the auspices of the Department of Defence (DoD) Digital Voice Processor Consortium. These test results were presented at the 1981 IEEE International Conference of Acoustics, Speech and Signal Processing [2]. Since then the voice processing algorithm has been refined, and a 16-kb/s mode has been incorporated. In addition, intelligibility and communicability tests were conducted at NRL on both the 9.6 and 16-kb/s modes. All of these recent developments are presented in this report.

According to these tests, the new voice processor operating at 9.6 kb/s provides a comparable speech quality to the presently deployed continuously variable slope delta (CVSD) modulator operating at 16 kb/s. Likewise, the new voice processor operating at 16 kb/s is comparable to CVSD operating at 32 kb/s.

Manuscript submitted May 11, 1982.

It is gratifying that this research effort to develop a new voice processor has made the transition into the development phase. The Navy is about to build 18 voice terminals utilizing the voice processor described in this report. They will be employed to test the operational flexibilities mentioned earlier.

BACKGROUND

Over the years numerous voice processors have been devised and deployed for operational use, such as: pulse code modulator (PCM) operating at 18.75 and 50 kb/s, CVSD at 16 and 32 kb/s, adaptive predictive coder (APC) at 6.4 and 9.6 kb/s, LPC at 2.4 kb/s, and channel vocoder at 2.4 kb/s. It is significant to note that a secure connection cannot be made between two different types of voice processors. They can communicate only through the regeneration of speech and redigitization which requires decryption of the speech data. Besides the loss of end-to-end encryption, this form of tandeming introduces speech degradations. For example, the diagnostic rhyme test (DRT) intelligibility score for a 16-kb/s CVSD is 93. When CVSD is tandemed with a 2.4-kb/s LPC, the overall intelligibility drops to 75 (15 points lower than the intelligibility of a 2.4-kb/s LPC operating by itself).

To eliminate these shortcomings, MRP generates several data rates (2.4, 9.6, and 16 kb/s) simultaneously. Speech data are so generated that the 16-kb/s mode utilizes the entire 9.6-kb/s data. Likewise, the 9.6-kb/s mode utilizes the entire 2.4-kb/s data except for the excitation parameters. Thus, the lower-rate data can be extracted directly from the higher-rate data. The embedded data structure makes direct rate-reduction possible by bit-stripping at a network node while maintaining end-to-end encryption. The 2.4-kb/s mode is directly interoperable with the 2.4-kb/s LPC currently under development by the DoD. The 9.6 or 16-kb/s modes provide higher speech quality for those users that have access to wideband channels.

Currently, the DoD Worldwide Digital System Architecture (WWDSA) study group is drafting recommendations for future DoD communication systems. One of the recommendations of this group is that future 16-kb/s terminals operating in tandem with other voice terminals must have an overall performance approximately equal to that of the weaker link. The MRP meets this requirement when the 16-kb/s mode operates with the DoD 2.4-kb/s LPC. Since the DoD 2.4-kb/s LPC is the only narrowband voice processor that will be deployed extensively, MRP meets the tandem performance requirement recommended by the WWDSA study group.

MRP as an Extension of the DoD 2.4-kb/s LPC

The voice processor has one important commonality; it uses the same speech synthesizer for all rates (see Fig. 1). In essence, the voice processing algorithm is a direct extension of the DoD 2.4-kb/s LPC. The difference between the 2.4-kb/s and the 9.6 or 16-kb/s rates is in the generation and transmission of the excitation signal for use in the synthesis filter. The DoD 2.4-kb/s LPC has as its excitation signal either a broadband signal that repeats quasi-periodically at the rate of the pitch frequency for voiced sounds, or random noise for unvoiced sounds. The excitation signal for the 9.6 or 16-kb/s mode is derived from the prediction residual signal, which is the ideal excitation signal for the LPC speech model.

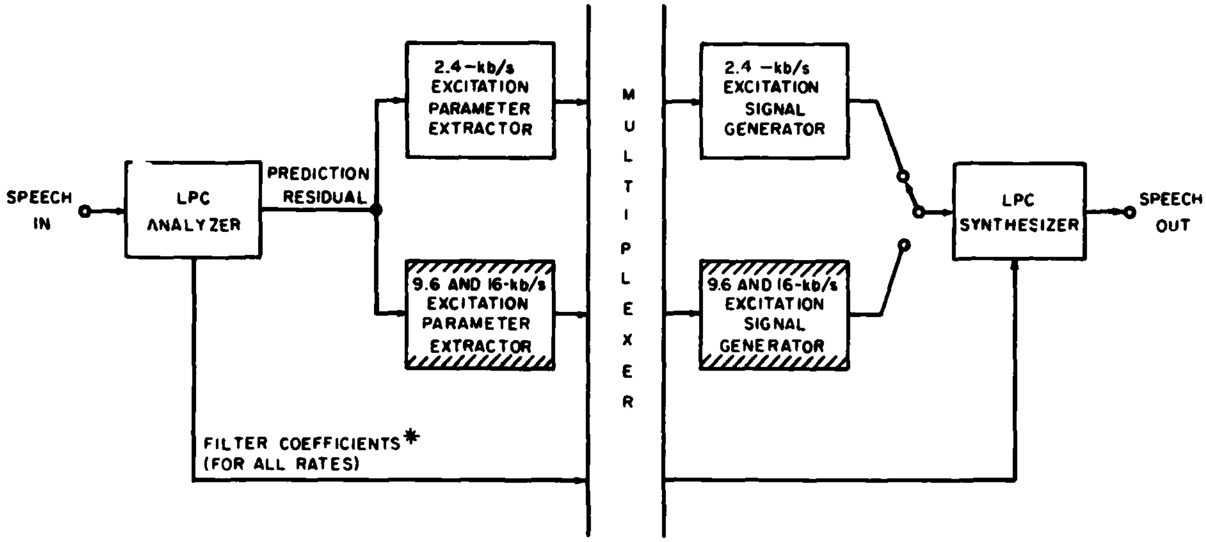

Fig. 1 — MRP voice processor. The hatched blocks are additions to the current 2.4-kb/s LPC device.

Wideband Residual-Excitation *vs* Baseband Residual-Excitation

The wideband residual-excited LPCs transmit residual samples for the entire residual passband. Each residual sample is quantized to one bit for the 9.6-kb/s rate or two bits for the 16-kb/s rate. The 9.6-kb/s wideband residual-excited LPC has been known for some time. This form of excitation produces speech that, in general, is raspy and fuzzy with background quantization noise plainly audible. Some listeners are not bothered by this speech quality whereas others actually prefer the 2.4-kb/s LPC.

The baseband residual-excited LPCs transmit residual samples within a low-frequency band (i.e., baseband) and regenerate the upperband at the receiver. The residual bandwidth is typically 1 kHz for the 9.6-kb/s rate and 2 kHz for the 16-kb/s rate with each residual sample quantized with as much as 3 bits. The baseband residual-excitation produces high quality speech which indicates that the human ear is somewhat tolerant to upper-frequency distortions if lower frequencies are well defined.

Residual Time-Sample Coding *vs* Residual Spectrum Coding

Once the baseband residual-excitation method is selected, there are still two possible ways to transmit residual information. The conventional method downsamples the low-pass filtered residual samples prior to encoding them [3,4]. Since MRP needs two down-sampling rates and the 9.6-kb/s data must be embedded in the 16-kb/s data, this approach does not lend itself to MRP implementation.

Thus, the residual spectrum coding method is selected for the MRP. A drawback of this approach is the need for a spectral conversion process (i.e., fast Fourier transform (FFT)). But the advantages, listed below, outweigh the disadvantages:

- Low-pass filtering and down sampling are not required.

- Extreme low-frequency components, not essential to speech communications, can be omitted from encoding to save bits.

*The details of this area will be discussed later in the text, but additional data to be included here are: error-correction codes for the 9.6 and 16-kb/s modes, and the unvoiced state of the 2.4-kb/s mode.

- The data rate can be changed at a small increment equal to the transmission rate of each spectral component (i.e., 6 bits/frame or 250 bits/s, as will be discussed later).

- The overhead data (sync bits, error protection bits, and side information bits) may be incremented at 6 bits/frame, which makes the bit-tradeoff between speech data and overhead data flexible.

- Residual spectral components for the 16-kb/s mode consist of those components for the 9.6-kb/s mode plus additional higher-frequency spectral components (an important aspect of the MRP implementation).

- Each spectral component at six bits allows for more error-resistant coding.

- The upperband can be regenerated at the receiver by simple spectral replication.

Upperband Regeneration

The baseband residual-excited LPC does not transmit upperband residual information. Hence, the receiver must produce an approximate upperband residual waveform from the received baseband information. The spectral envelope of the prediction residual is virtually flat due to inverse filtering. Thus, the upperband residual spectral envelope may be approximated by the baseband residual spectral envelope. If speech is unvoiced, the resulting approximation is satisfactory because the prediction residual is basically broadband random noise. Short-term amplitude variations are adequately reproduced by frame-to-frame updating of the baseband residual spectrum. On the other hand, if speech is voiced, the residual spectrum contains predominantly pitch harmonics under a flat spectral envelope. Since pitch harmonics are evenly spaced, the entire spectrum can also be reconstructed from the baseband spectrum.

The MRP regenerates the upperband spectrum by replication of the baseband spectrum. Advantages are: (i) it does not require much additional computation, and (ii) it does not distort the baseband spectrum. The disadvantage is that this produces nonuniformly spaced pitch harmonics. Since the baseband spectrum is not replicated at a multiple of the fundamental pitch-frequency, the composite spectrum is not expected to have evenly spaced pitch harmonics for voiced speech. The human ear is sensitive to this kind of pitch deformation. However, the unnatural tonal quality may be suppressed to an acceptable level by making the baseband bandwidth large enough (i.e., 1000 Hz in the 9.6-kb/s mode and 1917 Hz for the 16-kb/s mode). The human ear is somewhat deficient in crosscorrelating the upperband and the lowerband as demonstrated by coders of Sambur [5] and Watkins [6], where the output is a superposition of lowband speech and high-pass filtered narrowband speech. The upperband pitch-frequency is not only approximate, but it is also phase incoherent with that of the lowband. Nevertheless, both devices give satisfactory performance.

Bit Allocation

The 2.4-kb/s mode of the MRP must be interoperable with the DoD-standardized 2.4-kb/s LPC. This interoperable requirement determines the speech sampling rate of 8 kHz and the frame rate of 44.444 Hz (i.e., frame size of 180 samples). As a result, the number of bits available for the 2.4, 9.6, and 16-kb/s modes are 54, 216, and 360 bits/frame, respectively. The bit allocation for the DoD 2.4-kb/s mode, listed in Table 1, is firmly defined, which influences the bit utilization of the higher-rate modes.

As noted in Table 1, the 2.4-kb/s mode transmits ten filter coefficients if speech is voiced. If speech is unvoiced, however, it transmits only the first four filter coefficients with the 21 freed bits (one bit is unused) employed for error protection. The use of two different filter sizes, acceptable for the pitch-excited 2.4-kb/s LPC, is detrimental to the residual-excited 9.6 and 16-kb/s modes of the

Table 1 — DoD 2.4-kb/s LPC Design Parameters

GENERAL INFORMATION		
Speech sampling rate (kHz)		8
Frame rate (Hz)		44.444
Frame size (speech samples)		180
ENCODED DATA (bits/frame)		
Sync bit		1
Excitation parameters		
Amplitude		5
Pitch period		6
Voicing decision		1
Synthesis filter coefficients	(if voiced)	(if unvoiced)
Coefficient #1	5	5
2	5	5
3	5	5
4	5	5
5	4	0
6	4	0
7	4	0
8	4	0
9	3	0
10	2	0
Error-protection codes	0	20
Unused bit	0	1
		Total . . . 54 bits/frame

MRP. Changing the filter size at the voicing boundary alters the prediction residual characteristics, which in turn introduces flutter in the synthesized speech. Thus, the prediction residual must be generated by ten filter coefficients independent of the voicing decision. The filter coefficients and error-protection codes not transmitted by the 2.4-kb/s mode are included in the 9.6-kb/s mode.

Speech data for the 9.6-kb/s mode contain side information consisting of the fifth through the tenth filter coefficients (21 bits) if speech is unvoiced, or error-protection codes (21 bits of which one bit is not used) if speech is voiced. In addition, two redundant voicing decision bits are also included as side information to lessen the likelihood of the voicing decision being corrupted by transmission errors. At the higher-rate modes, all three voicing bits are subjected to a majority rule. Other transmitted parameters (i.e., spectral information) cannot be used as alternative voicing indicators because they are also susceptible to transmission errors. Also, they cannot be related reliably to the actual voicing decision made at the transmitter. These 23 bits are transmitted as part of the 9.6-kb/s data (which also will be used for the 16-kb/s mode).

Furthermore, two additional sync bits are included in both the 9.6 and 16-kb/s data. The total number of sync bits available for 9.6 and 16-kb/s modes are 3 and 5, respectively. A 72-bit up/down counter can be used to maintain synchronization for both higher-rate modes (note that the total number of bits per frame for the 9.6 and 16-kb/s modes are 216 and 360, respectively). Table 2 lists the allocation of bits for the higher-rate modes. The contents of the residual spectrum data will be presented in a later section.

Table 2 — Bit Allocation for MRP Voice Processor

2.4-kb/s mode (see Table 1)		54 bits/frame
9.6-kb/s mode		
All of the above	54	
Additional sync bits	2	
Side information	23	
Residual spectrum data (see Table 3)	137	
	Total . . . 216 bits/frame	
16-kb/s mode		
All of the above	216	
Additional sync bits	2	
Additional residual spectrum data (see Table 3)	142	
	Total . . . 360 bits/frame	

ALGORITHM DESCRIPTION

The basic principle of speech processing for the MRP is linear predictive analysis and synthesis. The analysis portion represents a future digitized speech sample $x(i)$ as a linear combination of past samples:

$$x(i) = \sum_{n=1}^{N} \alpha(n)x(i-n) + r(i) \qquad i = 1, 2, \ldots, I, \tag{1}$$

where $\alpha(n)$ is the n th prediction coefficient and $r(i)$ is the i th prediction residual sample. In terms of matrix notation, Eq. (1) may be denoted as

$$\underline{X} = \underline{AH} + \underline{R} \ . \tag{2}$$

The unbiased estimation of $\underline{A}$, by the application of the least squares method, is

$$\underline{A} = (\underline{H}^T\underline{H})^{-1}\,(\underline{H}^T\underline{X}) \ . \tag{3}$$

The solution of Eq. (3) has been well explored for implementing a 2.4-kb/s LPC. The MRP employs the method specified by the DoD-standardized 2.4-kb/s LPC [1,7] to derive and encode the filter coefficients. As stated earlier, the filter weights are common for all three data rates.

Based on the speech model of Eq. (1), speech is synthesized by

$$y(i) = \sum_{n=1}^{N} \hat{\alpha}(n)y(i-n) + e(i) \qquad i = 1, 2, \ldots, I, \tag{4}$$

where $y(i)$ is the synthesized speech sample, $\hat{\alpha}(n)$ is the quantized prediction coefficient, and $e(i)$ is the appropriate excitation signal determined by the data rate. For the 2.4-kb/s mode, the conventional excitation signal (random noise for unvoiced sounds or quasi-periodic broadband signal for voiced sounds) is used. For either the 9.6 or 16-kb/s mode, however, an approximate form of the prediction residual is transmitted. The remaining discussion is for the encoding of the prediction residual for the higher rates.

Residual Generation

For each frame, the prediction residual samples are generated by

$$r(i) = x(i) - \sum_{n=1}^{N} \alpha(n)x(i-n) \qquad i = 1, 2, \ldots, I. \tag{5}$$

The DoD-standardized 2.4-kb/s LPC determines the filter parameter encoding rule. The total number of filter weights is 10 (i.e., $N = 10$) independent of the voicing state. The total number of speech samples for each frame is 180 (i.e., $I = 180$).

Residual Spectrum Generation

Because of spectral quantization, the forward and inverse Fourier transforms tend to produce waveform discontinuities at the frame boundaries. Thus, it is necessary to overlap frames at the expense of transmission efficiency (i.e., it needs more bits to encode the same amount of residual information). Since wave form discontinuities can be heard as clicks, listening tests determined that an overlap size of 12 samples minimized this undesirable sound. A 12-sample overlap determines a Fourier transform size of 192, which can be implemented by a composite of six 32-point FFTs [8], or directly by the Winograd FFT [9]. A computationally simple trapezoidal window gives satisfactory results for this application. The 192 windowed prediction residual samples take the form:

$$\begin{aligned} r'(i) &= \left[\frac{i}{13}\right] r(i) && i = 1, 2, \ldots, 12, \\ r'(i) &= r(i) && i = 13, 14, \ldots, 180, \\ r'(i) &= \left[\frac{(193-i)}{13}\right] r(i) && i = 181, 182, \ldots, 192, \end{aligned} \tag{6}$$

where $r'(i)$ is the ith windowed and time-overlapped residual sample.

Since the residual samples are real and only a portion of the residual spectrum is transmitted, the use of a half-size complex Fourier transform is advantageous [10]. The 192 windowed prediction residual samples ($r'(i)$, $i = 1, 2, \ldots, 192$) are loaded alternately into the real and imaginary parts of a 96-word complex buffer (or simply treat every other windowed residual sample as being phase-shifted by $\pi/2$ radians). By the use of a specially generated 96-point complex fast Fourier transform (listed in the Appendix), the scrambled prediction residual spectrum is obtained. The resulting complex spectrum is of the form:

$$C(k) = A(k) + jB(k) \qquad k = 1, 2, \ldots, 96. \tag{7}$$

Since only the baseband spectral information is transmitted, a limited number of Fourier components (i.e., $k = 3$ to 47 as noted in Table 3) need be obtained by the following descrambling process [10]:

$$\begin{bmatrix} R(k) \\ X(k) \end{bmatrix} = \begin{bmatrix} A1 \\ B1 \end{bmatrix} + \begin{bmatrix} B1 & -A2 \\ -A2 & B1 \end{bmatrix} \begin{bmatrix} \mathrm{COS}(\theta(k)) \\ \mathrm{SIN}(\theta(k)) \end{bmatrix} \qquad k = 3, 4, \ldots, 47, \tag{8}$$

where

$$\begin{aligned} A1 &= A(k) + A(98-k), \\ A2 &= A(k) - A(98-k), \\ B1 &= B(k) + B(98-k), \\ B2 &= B(k) - B(98-k), \\ \theta(k) &= \pi(k-1)/96, \end{aligned}$$

and $R(k)$ and $X(k)$ are, respectively, the real and imaginary components of the windowed prediction residual spectral components. The maximum amplitude spectral component in the baseband of the 9.6-kb/s mode (5th through the 25th spectral indices) is transmitted as a frame-to-frame amplitude normalization factor. The encoding method for the individual spectral components will be discussed later.

Upperband Residual Regeneration at the Receiver

The upperband residual spectrum is obtained by replicating the baseband residual spectrum. For the 9.6-kb/s mode,

$$\begin{aligned} R'(k+21i) &= R'(k) \qquad & i &= 1,\ 2,\ 3, \\ X'(k+21i) &= X'(k) \qquad & k &= 5,\ 6,\ \ldots,\ 25, \end{aligned} \tag{9}$$

where the $R'(k)$ and $X'(k)$ are, respectively, the kth real and imaginary parts of the amplitude-weighted baseband residual spectral components. For the 16-kb/s mode,

$$\begin{aligned} R'(k+45) &= R'(k) \qquad & k &= 3,\ 4,\ \ldots,\ 47, \\ X'(k+45) &= X'(k). \end{aligned} \tag{10}$$

The resulting 96 complex spectral components are converted to 192 time-samples by the inverse Fourier transform (see Appendix). The 12 leading time-samples of the current frame are overlapped with the 12 trailing time-samples of the preceding frame. The resulting time-samples are the excitation signal. Thus,

$$\begin{aligned} e(j,1) &= e'(j-1,\ 181) + e'(j,\ 1), \\ e(j,2) &= e'(j-1,\ 182) + e'(j,\ 2), \\ &\ \ \vdots \\ e(j,\ 12) &= e'(j-1,\ 192) + e'(j,\ 12), \\ e(j,\ 13) &= e'(j,\ 13), \\ &\ \ \vdots \\ e(j,\ 180) &= e'(j,\ 180), \end{aligned} \tag{11}$$

where $e'(j,k)$ and $e(j,k)$ are, respectively, the kth time-samples of the jth frame before and after time-overlapping ($k = 1,\ 2,\ \ldots,\ 180$). These samples are fed into the speech synthesizer, which is based on Eq. (4).

RESIDUAL SPECTRUM ENCODING

Since MRP utilizes the same filter coefficients for all rates in the synthesis filter, the performance of MRP at the different rates is dependent on the quality of the excitation signal driving the synthesis filter. Since the excitation signal is obtained from the prediction residual, the residual coding is a critical element in the MRP design. For reasons mentioned earlier, the higher-rate modes of the MRP transmit baseband residual information in terms of spectral components. The selection of a particular residual spectrum encoding method is based on the considerations listed in the rest of this section.

Design Objectives

1. High-quality speech: MRP is designed to provide operational flexibility in voice communication by embedding lower-rate data in higher-rate data. However, MRP will not be widely accepted unless it is capable of producing speech quality comparable to other single-rate processors operating at similar data rates. Since baseband information is critical to speech quality, it is quantized with a finer resolution.

2. Simpler implementation: The Navy is in the process of procuring 18 MRP terminals for deployment at selected communication centers. Therefore, a computationally efficient spectrum encoding method is desired. The computation time required by MRP is approximately 30% above that required by the DoD 2.4-kb/s LPC. A quality improvement gained by a more complicated residual encoding scheme must be weighed against the resulting impact on hardware complexity.

3. Nonspeech signal processing: The commercial telephone was originally designed for the transmission of analog speech signals. Recently, the telephone is increasingly becoming a means to transmit nonspeech signals. Similarly, once the MRP is deployed, the 16-kb/s mode may be used for transmitting nonspeech signals such as facsimile, graphics, and other information within a limited bandwidth. The ability of MRP to transmit nonspeech signals can be a desirable feature, particularly in an emergency. Thus, residual encoding methods highly customized for speech signals (viz., the use of a long-term predictor based on the fundamental pitch-period) have been avoided in the MRP.

Baseband Bandwidth

The choice of a baseband bandwidth to achieve the highest speech quality is a tradeoff between the number of bits available to encode the residual information and the number of bits assigned to each residual component. The baseband bandwidth has been typically around 1 kHz for residual-excited LPCs operating at 9.6 kb/s [1, 3, 4].

The number of bits available to encode the residual information at the 9.6-kb/s mode is 137 bits/frame (see Table 2). Experimentation indicates that the generation of high-quality speech needs six bits for each complex spectral component. Thus, the 9.6-kb/s mode of the MRP can transmit 21 spectral components (i.e., 5th through the 25th). Since each spectral component is separated by 41.67 Hz, the baseband bandwidth for the 9.6-kb/s mode is from 167 Hz to 1,000 Hz.

On the other hand, the number of bits available for encoding the residual at the 16-kb/s mode is 279 bits/frame (see Table 2). Hence, the 16-kb/s mode can transmit 45 spectral components (i.e., 3rd through the 47th). Therefore, the baseband bandwidth for the 16-kb/s mode is from 83 Hz to 1917 Hz as indicated in Table 3.

Amplitude Normalization Factor

The maximum spectral component within the 9.6-kb/s baseband is used as a frame-to-frame amplitude normalization factor for the 9.6-kb/s mode. This component is quantized semi-logarithmically to 6 bits. Three additional bits are included in the 9.6-kb/s mode for error protection of this parameter. The 16-kb/s mode uses the 9.6-kb/s normalization factor without any additional modification.

All amplitude spectral components are scaled by this factor prior to encoding. Since the maximum amplitude spectral factor is taken from the 9.6-kb/s baseband, some of the normalized amplitude spectral components may need to be clamped. Hereafter, unless stated otherwise, the amplitude spectral component is referred to as the normalized amplitude spectral component whose magnitude lies between 0.0 and 1.0.

Residual Spectrum Characteristics

The residual amplitude spectrum has a recognizable structure. If the input speech waveform is unvoiced (an example of this case is shown in Fig. 2), the resulting residual amplitude spectrum is relatively flat. If speech is voiced (an example of this case is shown in Fig. 3), the residual amplitude spectrum has pitch harmonics under a relatively flat spectral envelope. The pitch harmonics, however, are rather irregular even for high-pitch female voices because the frequency resolution is a coarse 41.67 Hz. Thus, intraframe amplitude spectrum correlation cannot be readily exploited to save transmission bits.

Table 3 — Residual Encoding Information

GENERAL INFORMATION	
Total bandwidth (Hz)	4000
Complex frequency components (samples)	96
Frequency component spacing (Hz)	41.67
Baseband spectral indices	
9.6-kb/s mode	5-25
16-kb/s mode	3-47
Baseband bandwidth (Hz)	
9.6-kb/s mode	167-1000
16-kb/s mode	83-1917
ENCODED RESIDUAL DATA (bits/frame)	
9.6-kb/s mode	
Maximum amplitude spectral component	6
Error protection for maximum amplitude	3
21 complex frequency components	126
Unused bits	2
	Total . . . 137
16-kb/s mode	
All of the above (embedded)	137
24 additional complex frequency components (with 2 unused bits in the 9.6-kb/s mode)	142

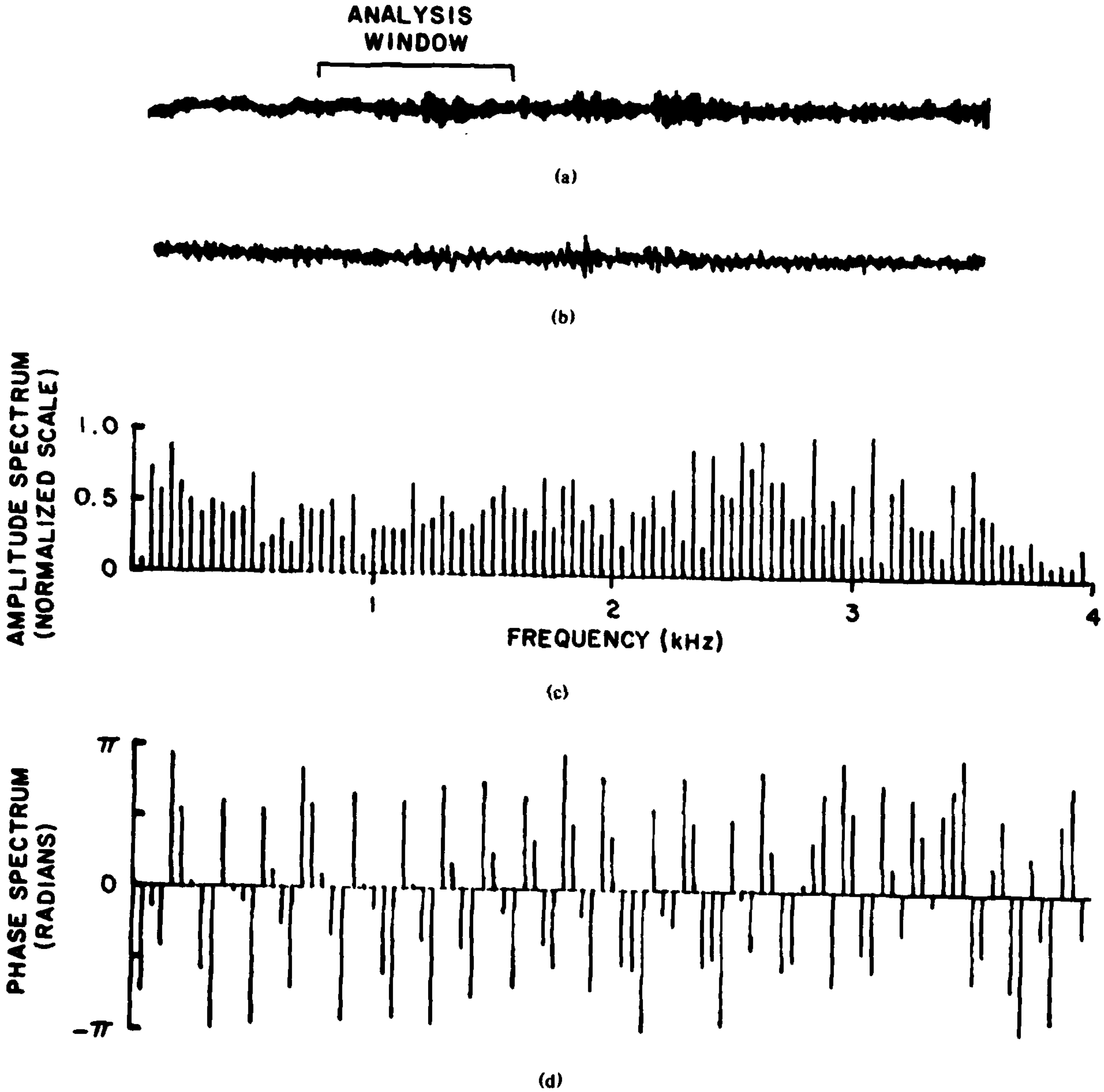

Fig. 2 — Speech waveform, prediction residual waveform and prediction residual spectrum (voiced case)

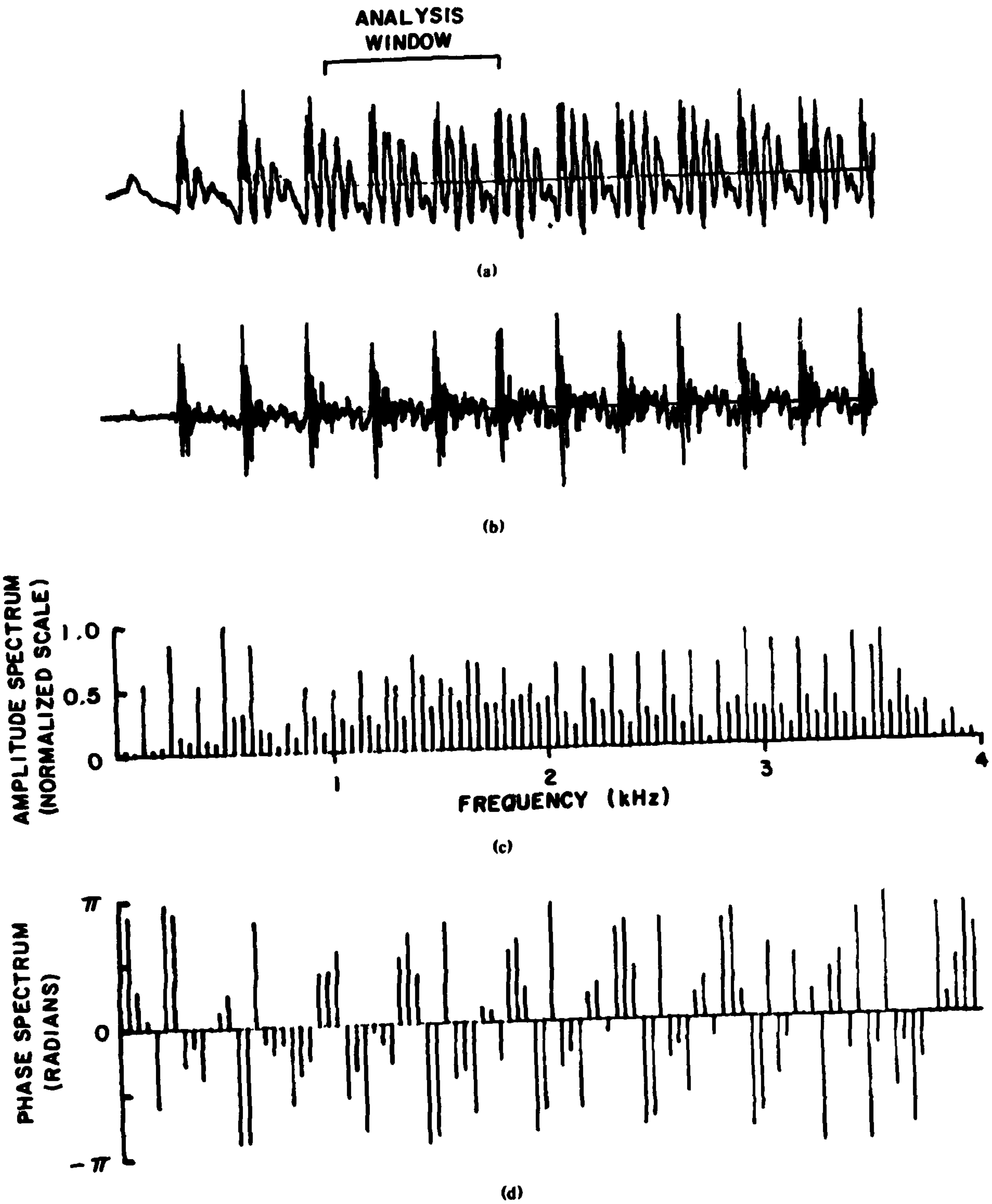

Fig. 3 — Speech waveform, prediction residual waveform and prediction residual spectrum (unvoiced case)

On the other hand, each residual phase spectral component is random from one frame to the next. This is true even when speech is voiced because the LPC frame is completely asynchronous with respect to the pitch cycle. Also, residual phase spectral components a pitch-frequency apart do not show any degree of correlation, unlike the residual amplitude spectral components.

Residual Spectrum Encoder

A simple residual spectrum encoder is one that encodes each amplitude and phase spectral component independently, as in the 9.6-kb/s rate of the MRP previously published [1]. In that report, each phase spectral component is quantized to three bits and each amplitude spectral component is quantized to two bits. The phase component has a finer resolution than the amplitude component because it carries timing information, which is an important aspect of the excitation signal.

Because the previous approach lacked bit-tradeoff flexibility between amplitude and phase components of the same frequency, an alternative residual spectrum quantizer (but computationally more demanding) is described in the rest of this section.

Encoding each amplitude and phase spectral component of the same frequency jointly (i.e., block encoding) has the following three advantages:

1. Flexible quantization: The number of quantization steps for either the amplitude or phase spectral component can be made an integer number rather than a binary number as in the previous approach. Thus, the phase resolution can be traded with the amplitude resolution more easily to achieve better quality speech.

2. Amplitude-dependent phase resolution: The ear is more sensitive to stronger amplitude components. Thus, phase spectral components with smaller amplitude spectral values (i.e., approximately −15 dB or less with respect to the peak value) can be quantized with a coarser step without introducing audible distortions in the synthesized speech. Amplitude-dependent phase quantization is feasible with block coding.

3. Introduction of diversified phase angles: Experimentation indicates that synthesized speech sounds more natural if the decoded phase spectral components are diversified. In the block-coding approach, each amplitude quantization level is associated with a phase level. Thus, the decoded phase spectral components of the block-coding approach have more phase angles than can be realized by the previous approach.

Since the phase and amplitude spectra are uncorrelated, the phase and amplitude quantizers may be designed separately. The quantizer was designed by going through the following steps:

1. Assignment of bits for each complex spectral component: For the generation of near toll-quality speech, MRP needs six bits to encode each complex spectral component. The following residual spectrum encoder was designed on that basis.

2. The number of quantization steps needed for the phase spectral component: The phase spectrum defines how each spectral component is phased in reference to the beginning of the LPC frame. The required phase resolution can be determined only through extensive listening tests because no measurement exists to tell us how the ear processes phase information. According to listening tests with various phase resolutions, a 10-level phase quantization generated high quality speech.

3. The number of quantization steps for amplitude spectral components: Since each complex spectral component is represented by a 6-bit word (i.e., 64 possible combinations), this allows seven amplitude quantization steps with the phase resolution coarser when the amplitude spectral component is small (i.e., approximately −15 dB or less). Thus, a seven-level amplitude quantization is possible.

4. Quantization step size for phase spectral components: Because the phase spectral component is uniformly distributed, uniform quantization steps can be used for the phase information. To achieve a diversified phase angle, the initial phase quantization level associated with an amplitude quantization level is alternately staggered.

5. Quantization step size for amplitude spectral components: A seven-amplitude quantizer was designed from the probability density function of the amplitude spectral components shown in Fig. 4. This curve was obtained from 1 600 000 amplitude spectral components from both male and female voices. A 7-level quantizer is based on the amplitude transfer characteristic:

$$\begin{aligned} y(x) &= x_1/2, && \text{if } 0 \leqslant x \leqslant x_1, \\ &= (x_1 + x_2)/2, && \text{if } x_1 < x \leqslant x_2, \\ &= (x_2 + x_3)/2, && \text{if } x_2 < x \leqslant x_3, \\ &= (x_3 + x_4)/2, && \text{if } x_3 < x \leqslant x_4, \\ &= (x_4 + x_5)/2, && \text{if } x_4 < x \leqslant x_5, \\ &= (x_5 + x_6)/2, && \text{if } x_5 < x \leqslant x_6, \\ &= (x_6 + 1)/2, && \text{if } x_6 < x \leqslant 1, \end{aligned} \tag{12}$$

where x is the normalized input amplitude, $y(x)$ is the output amplitude, and $x_1, x_2, \ldots, x_6$ are input amplitude break points.

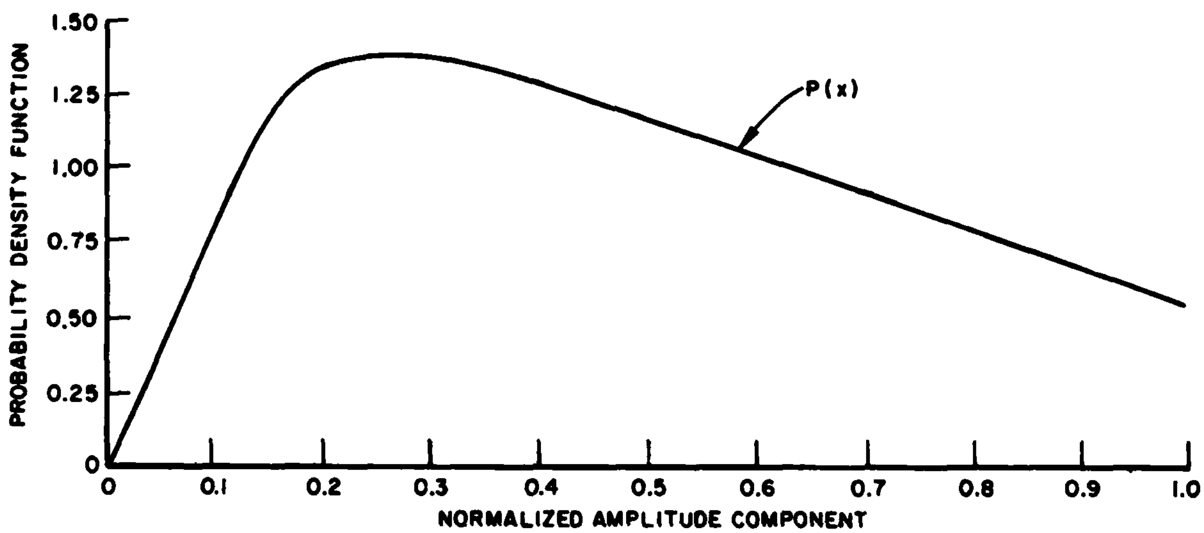

Fig. 4 — Probability density function of residual amplitude spectral components (peak amplitude normalized)

The quantization error is defined as the quantized output amplitude minus the input amplitude:

$$\epsilon(x) = y(x) - x. \tag{13}$$

The mean-square value of the quantization error is

$$\overline{\epsilon^2} = \sum_{x=0}^{x_1} (y(x) - x)^2 \, p(x) + \sum_{x=x_1}^{x_2} (y(x) - x)^2 \, p(x) + \ldots + \sum_{x=x_6}^{1} (y(x) - x)^2 \, p(x). \tag{14}$$

The quantizer parameters $(x_1, x_2, \ldots, x_6)$ which minimize the above error have been computed. The resulting quantizer has the amplitude transfer characteristic:

$$
\begin{aligned}
y(x) &= 0.078, \text{ if } 0.0 \leqslant x \leqslant 0.156 \\
&= 0.219, \text{ if } 0.156 < x \leqslant 0.281 \\
&= 0.344, \text{ if } 0.281 < x \leqslant 0.406 \\
&= 0.469, \text{ if } 0.406 < x \leqslant 0.531 \\
&= 0.609, \text{ if } 0.531 < x \leqslant 0.688 \\
&= 0.766, \text{ if } 0.688 < x \leqslant 0.844 \\
&= 0.922, \text{ if } 0.844 < x \leqslant 1.0
\end{aligned} \tag{15}
$$

Table 4 lists the 64 possible values for each encoded complex spectral component. As already mentioned, there are seven amplitude quantization steps. For each amplitude value there are ten possible phase values except for the two lowest amplitude levels which have eight and six phase values respectively. Since the amplitude quantization steps are nearly equal, MRP can encode non-speech signals confined within the baseband.

The speech bit stream has unequal sensitivity to transmission errors because of modem characteristics. The MRP performance can be made more robust if the most sensitive bits describing the complex spectral component are assigned to the most error resistant bits of the modem. One approach is to divide the unit circle (where the 64 points listed in Table 4 are located) into four quadrants. The 16 points in each quadrant are identified by 2 bits, i.e., b_1 and b_2 in a six-bit word denoted by $(b_1, b_2, b_3, b_4, b_5, b_6)$. Likewise, each quadrant is further divided into four sectors, and the four points in each sector are identified by b_3 and b_4. Finally, each sector is divided into four sections, and the four points in each section are identified by b_5 and b_6.

Bits b_1 and b_2 are the most sensitive because a one-bit error makes a decoded complex spectral component in one of the adjacent quadrants (but not in the opposite quadrant). Thus, b_1 and b_2 should be mapped into the most error resistant bits of the modem. On the other hand, b_5 and b_6 are the least significant bits.

REAL TIME SIMULATION OF 9.6 AND 16-kb/s RATES

The baseband residual excited LPC is simulated for real time operation on the MVP. MVP was originally built to demonstrate in real time the DoD 2.4-kb/s LPC algorithm to be used in the Advanced Narrowband Digital Voice Terminal. The existing 2.4-kb/s software was altered to accomodate the excitation signals of the 9.6 and 16-kb/s rates, but otherwise remained the same.

MVP Hardware Description

The MVP was built primarily for signal processing algorithms operating in a real-time environment. The MVP has a parallel architecture which allows arithmetic processing, address generation, multiplication, logic testing, and branching all within the same instruction cycle (350 ns). Some of the features of the machine are:

- two arithmetic logic units (ALUs)—one for data processing, and the other for address generation,
- two memorys—6144 words (70 bits per word) of program memory, and 6094 words (16 bits per word) of data memory,
- input/output through modem and teletype,

Table 4 — Quantization Table for Each Complex Spectral Component

Index	Amplitude	Phase (deg.)
1	0.922	18
2	0.922	54
3	0.922	90
4	0.922	126
5	0.922	162
6	0.922	198
7	0.922	234
8	0.922	270
9	0.922	306
10	0.922	342
11	0.766	36
12	0.766	72
13	0.766	108
14	0.766	144
15	0.766	180
16	0.766	216
17	0.766	252
18	0.766	288
19	0.766	324
20	0.766	0
21	0.609	18
22	0.609	54
23	0.609	90
24	0.609	126
25	0.609	162
26	0.609	198
27	0.609	234
28	0.609	270
29	0.609	306
30	0.609	342
31	0.469	36
32	0.469	72
33	0.469	108
34	0.469	144
35	0.469	180
36	0.469	216
37	0.469	252
38	0.469	288
39	0.469	324
40	0.469	0
41	0.344	18
42	0.344	54
43	0.344	90
44	0.344	126
45	0.344	162
46	0.344	198
47	0.344	234
48	0.344	270
49	0.344	306
50	0.344	342
51	0.219	45
52	0.219	90
53	0.219	135
54	0.219	180
55	0.219	225
56	0.219	270
57	0.219	315
58	0.219	0
59	0.078	30
60	0.078	90
61	0.078	150
62	0.078	210
63	0.078	270
64	0.078	330

- sixteen vectored interrupts,
- two's complement fractional arithmetic,
- two programmable sample rate counters,
- two analog-to-digital and two digital-to-analog converters,
- program loading via card reader,
- user interface via front panel or teletype.

Software Description

The implementation of the higher rates on the MVP was constrained by the available computation time. This limiting factor necessitated that the quantization of the residual spectral values be done directly on the real and imaginary values coming out of the FFT rather than the more computational demanding quantization of their phase and amplitude components. Otherwise the real-time simulation closely follows the algorithmic description of the previous section. The block diagram of the main subroutines used for generation of the 9.6 and 16-kb/s excitation signals are shown in Fig. 5, and Table 5 lists the computation times for each subroutine.

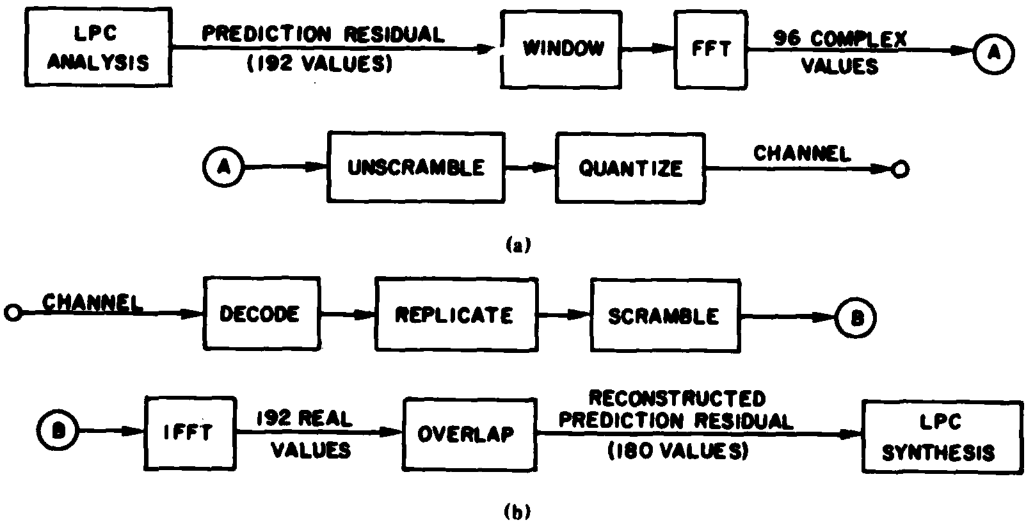

Fig. 5 — Block diagram of subroutines used to generate 9.6 and 16-kb/s excitation signal

Table 5 — Execution Time for Subroutines Running in the MVP

Subroutine	Time (ms)	
	9.6 kb/s	16 kb/s
Window	0.04	0.04
FFT	2.25	2.25
Unscramble	0.19	0.33
Quantize	0.22	0.35
Total	2.70	2.97

(a) Transmitter

Subroutine	Time (ms)	
	9.6 kb/s	16 kb/s
Decode	0.13	0.22
Replicate	0.20	0.14
Scramble	0.73	0.73
IFFT	2.35	2.35
Overlap	0.03	0.03
Total	3.44	3.47

(b) Receiver

TEST RESULTS

Diagnostic Rhyme Test

Quantitative evaluations of synthesized speech can be made by means of the DRT. The DRT word list comprises 448 monosyllable rhyming word pairs in which initial consonants differ by only a single feature. An important objective of the DRT [11] is to determine speech perception as influenced by process parameters (the parameter update rate, the number of bits for each parameter, and the choice of parameters). The test provides a measure of intelligibility and allows one to evaluate the discriminability of six distinctive features: voicing, nasality, sustention, sibilation, graveness, and compactness.

The DRT results for the 9.6 and 16-kb/s rates presented in this section were obtained from the real-time simulation of the algorithm on the MVP as described in the previous section. It is useful to compare the 9.6 and 16-kb/s MRP with the 16 and 32-kb/s CVSD respectively. Table 6 shows that for the back-to-back mode the 9.6-kb/s MRP scores one point lower than the 16-kb/s CVSD and the 16-kb/s MRP scores one point lower than the 32-kb/s CVSD. With acoustic background noise interference, the 9.6-kb/s MRP performance is comparable to the 16-kb/s CVSD. Unfortunately test results for the 32-kb/s CVSD with background noise are not available.

As noted from Table 6, DRT scores do not differ significantly among higher-rate processors, meaning that they all have acceptably good initial consonant intelligibility. For these processors, communicability tests are more meaningful and scores will be presented in the next section.

Table 6 — Comparison of DRT Scores of 9.6 and 16-kb/s MRP with 16 and 32-kb/s CVSD

TEST CONDITIONS	Noise Level (dB)*	No. of Spkrs	DRT Scores			
			MRP 9.6	CVSD 16	MRP 16	CVSD 32
Back-to-back mode	—	3M	93	94	95	96
WITH ACOUSTIC BACKGROUND NOISE						
Shipboard noise	82	3M	89	93	89	—
E3A noise	87	3M	86	89	91	—
Tank noise	112	3M	90	87	91	—
AVERAGE			90	91	92	—

*The normal speaking level is approximately 115 dB.

Table 7 lists comparisons between 9.6-kb/s MRP and 16-kb/s CVSD for a greater number of test conditions than were available for Table 6. The average score over all the conditions between the two processors is nearly the same. The DRT scores included in Table 7 were obtained from independent testing done in 1980 by the DoD Digital Voice Processor Consortium. At that time, the 16-kb/s mode of the MRP was not implemented for real-time operation. Thus, no scores are presented here for that rate.

The intelligibilities of both the 9.6-kb/s MRP and the 16-kb/s CVSD are not impaired by errors as much as 1%. The error performance between these processors, however, cannot be compared directly because of the difference in data rates. If the error is 5% at 16 kb/s for a given channel, the error rate at 9.6 kb/s is surely less for the same channel. If the error rate is as much as 5% at 9.6 kb/s, the MRP has an option to use the 2.4-kb/s mode.

Table 7 — Comparison of DRT Scores of 9.6-kb/s MRP with 16-kb/s CVSD

TEST CONDITIONS	Noise Level (dB)*	No. of Spkrs	DRT Scores	
			MRP 9.6	CVSD 16
Back-to-back mode	—	9M/9F	90	93
WITH ACOUSTIC BACKGROUND NOISE				
Office noise	63	3M/3F	88	90
Airborne command post noise	85	3M/3F	85	86
Shipboard noise	82	3M/3F	84	85
Helicopter noise	125	3M/3F	64	70
E3A noise	87	3M/3F	90	91
P3C turbo prop noise	105	3M/3F	86	85
Destroyer noise	78	3M/3F	78	76
Helicopter carrier noise	76	3M/3F	87	83
Jeep noise	92	3M/3F	87	84
Tank noise	112	3M/3F	86	83
WITH TRANSMISSION ERROR				
0.5%	—	3M/3F	89	90
1.0%	—	3M/3F	89	90
2.0%	—	3M/3F	85	87
5.0%	—	3M/3F	75	85
UNDER TANDEM ARRANGEMENT				
Self tandem	—	3M/3F	86	87
Output into 2.4-kb/s LPC	—	3M/3F	79	75
Input from 2.4-kb/s LPC	—	3M/3F	79	82
AVERAGE			84	85

*The normal speaking level is approximately 115 dB.

NRL Communicability Test

While the DRT is an excellent tool for testing the initial consonant, it is not intended to examine user's subjective opinions of communicability. A conversational test using live two-way communication to measure usability of voice systems was developed at NRL by Schmidt-Nielsen and S. Everett [12]. The NRL Communicability Test is the name given to the test. The NRL test uses two participants at a time with a communication task similar to the pencil-and-paper game "battleship". In this game, players place "ships" on a grid and then attempt to sink one another's ships by taking turns "shooting" at specified squares on the grid. There are four rating scales to be filled out after the game is completed. Figure 6 shows the average score (indicated by ∇) for several processors using six amateur radio operators as participants. Each of the six radio operators tested each processor three times. The horizontal line represents the range of the standard deviation around its mean.

CONCLUSION

The effectiveness of a voice communication system cannot be evaluated solely on the basis of speech intelligibility and quality. A high-rate system that is capable of providing acceptable communicability can become inoperative if the network is overloaded or disrupted by natural or man-made interference. The communication system must be designed so as to survive in the event of an emergency.

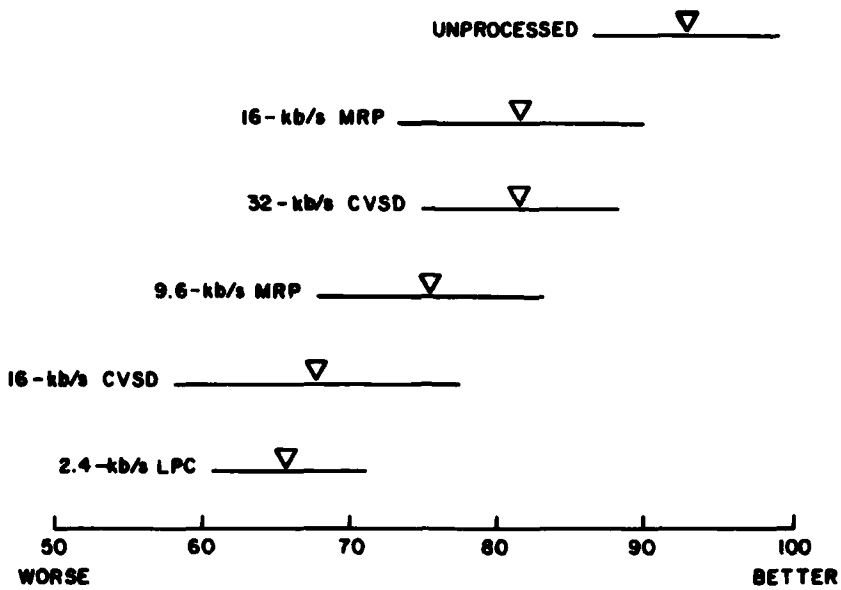

Fig. 6 — NRL Communicability Test results using amateur radio operators as participants

The MRP, as presented in this report, is designed to provide operational flexibility in voice communication by integrating narrowband and wideband resources into a single capability. It utilizes a single voice processing principle to generate both high and low data rates simultaneously. Since lower-rate data are embedded in the higher-rate data, a direct rate-conversion is possible by bit stripping at a network node as necessary. The lowest data rate of the MRP is 2.4 kb/s and it is directly interoperable with other 2.4-kb/s voice processors being developed by DoD. The selected higher data rates are 9.6 and 16 kb/s, but any other rate above 9.6 kb/s can be realized.

In addition to operational flexibilities and simplified hardware logistics, MRP is capable of providing speech quality that is comparable to other fixed-rate voice processors at similar data rates.

In conclusion, the ultimate objective of a voice terminal is to provide a reliable, survivable, and robust performance under all operational conditions, particularly in an emergency. The MRP is a step toward reaching this objective.

ACKNOWLEDGMENTS

MRP was initially developed and tested using our fine computer facility which is maintained in good working condition by Ruth Phillips of NRL. The communicability tests were made possible by the technical support provided by Howard Murphy of NRL.

The authors express their appreciation to TRW members, Drs. Zibman, Kau, and Tsou who were responsible for the initial MRP software implementation on the MVP. Finally the authors gratefully acknowledge the support of Bob Martin of the Naval Electronic Systems Command.

REFERENCES

1. G. Kang, L. Fransen, and E. Kline, "Multirate Processor (MRP) for Digital Voice Communications," NRL Report 8295, Mar. 1979.

2. G. Kang, L. Fransen, and E. Kline, "Mediumband Speech Processor with Baseband Residual Spectrum Encoding," Record of the 1981 IEEE ICASSP, p. 820-823.

3. R. Viswanathan, A. Higgins, W. Russell, and J. Makhoul, "Baseband LPC Coder for Speech Transmission Over 9.6 kb/s Noisy Channels," ICASSP-80, p. 348-351.

4. M. Dankberg and D. Wong, "Development of a 4.8—9.6 kbps RELP Vocoder," ICASSP-79, p. 554—557.

5. M. R. Sambur, "High-Quality 9.6 kb/s Algorithm that Satisfies the Embedded Bit Concept," Internatonal Conference on Communications, 1, paper 12.A.2.1-4, IEEE, June 1978.

6. H.E. Watkins, "Description of a Hybrid 7.2 kbps Vocoder," ICASSP-79, p. 546-549.

7. T. Tremain, J. Fussell, R. Dean, B. Abzug, M. Cowing, and P. Boudra, Jr., "Implementation of Two Real Time Narrowband Speech Algorithms," EASCON '78 Record, p. 698-708.

8. A. Oppenheim and R. Schafer, *Digital Signal Processing*, Prentice Hall, Inc., Englewood Cliffs, New Jersey, 1975.

9. H. Silverman, "An Introduction to Programming the Winograd Fourier Transform Algorithm (WFTA)," IEEE Transactions on Acoustics, Speech and Signal Processing, **ASSP-25**, 2, April 1977.

10. E. Brigham, *The Fast Fourier Transform*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1974.

11. W. Voiers, "Diagnostic Evaluation of Speech Intelligibility", in *Speech Intelligibility and Speaker Recognition*, M.E. Hawley (ed.), Dowden, Hutchinson and Ross, Stroudsburg, Pa., 1977.

12. A. Schmidt-Nielsen and S. Everett, "A Conversational Test for Comparing Voice Systems Using Working Two-Way Communication Links," NRL Report 8583, June 1982.

Appendix

TIME-TO-FREQUENCY AND FREQUENCY-TO-TIME CONVERSIONS USING 96-POINT COMPLEX FFT ALGORITHM

A 96-point complex FFT FORTRAN program was developed by NRL for the MRP implementation. This algorithm has been programmed for the NRL-owned special signal processor for the 9.6 and 16-kb/s modes of the MRP.

To enhance numerical accuracy through the FFT, a division by two is incorporated at each summing point of two vectors. As a result, the overall gain through the forward and inverse Fourier transforms is 0.375.

```
C
C       ********* 96-POINT COMPLEX FFT/IFFT SUBROUTINE ********
C
C
C
C
        SUBROUTINE FOURT(DATA,ISIGN)
        DIMENSION DATA(1),WORK(192)
        DIMENSION INDX1(96)
        DIMENSION ASIN(4),ACOS(4),BSIN(32),BCOS(32)
        DATA INDX1/1,97,49,145,25,121,73,169,13,109,61,157,37,133,85,181
        1,7,103,55,151,31,127,79,175,19,115,67,163,43,139,91,187,3,99,51
        1,147,27,123,75,171,15,111,63,159,39,135,87,183,9,105,57,153,33,129
        1,81,177,21,117,69,165,45,141,93,189,5,101,53,149,29,125,77,173,17
        1,113,65,161,41,137,89,185,11,107,59,155,35,131,83,179,23,119,71
        1,167,47,143,95,191/
        DATA ACOS/.7071068,.9238795,.9807853,.8314696/
        DATA ASIN/-.7071068,-.3826835,-.1950903,-.5555702/
        DATA BCOS/1.0,.9978589,.9914448,.9807853,.9659258,.9469302
        1,.9238795,.8968728,.8660254,.8314696,.7933533,.7518398,.7071068
        1,.6593458,.6087614,.5555702,.5000000,.4422887,.3826834,.3214395
        1,.2588190,.1950902,.1305261,.0654032,.0000000,-.0654032
        1,-.1305262,-.1950904,-.2588191,-.3214394,-.3826834,-.4422887/
        DATA BSIN/.0000000,-.0654031,-.1305262,-.1950903,-.2588190
        1,-.3214395,-.3826835,-.4422887,-.5000000,-.5555702,-.6087615
        1,-.6593459,-.7071068,-.7518398,-.7933533,-.8314696,-.8660254
        1,-.8968728,-.9238796,-.9469301,-.9659259,-.9807853,-.9914449
        1,-.9978589,-1.0000000,-.9978589,-.9914448,-.9807853,-.9659259
        1,-.9469301,-.9238796,-.8968728/
        RTHLF=.7071067812
```

```
C
C     SHUFFLE DATA BY DIGIT REVERSAL FOR GENERAL N
C
        J=1
        DO 260 I=1,96
        IND=INDX1(I)
        WORK(J)=DATA(IND)
        WORK(J+1)=DATA(IND+1)
260     J=J+2
        DO 270 I=1,192
270     DATA(I)=WORK(I)
C
C     MAIN LOOP FOR FACTORS OF TWO.
C     W=EXP(ISIGN*2*PI*SQRT(-1)*M/(4*MMAX)).  CHECK FOR W=ISIGN*SQRT(-1)
C     AND REPEAT FOR W=W*(1+ISIGN*SQRT(-1))/SQRT(2).
C
        IPAR=2
      DO 340 K1=1,192,4
      TEMPR=DATA(K1+2)/2
      TEMPI=DATA(K1+3)/2
      DATA(K1+2)=DATA(K1)/2-TEMPR
      DATA(K1+3)=DATA(K1+1)/2-TEMPI
      DATA(K1)=DATA(K1)/2+TEMPR
340   DATA(K1+1)=DATA(K1+1)/2+TEMPI
        IPAR=2
      MMAX=2
        L123=0
360   IF(MMAX-32)370,600,600
370   LMAX=MAX0(4,MMAX/2)
      DO 570 L=2,LMAX,4
      M=L
      IF(MMAX-2)420,420,380
380     L123=L123+1
        WR=ACOS(L123)
        WI=ASIN(L123)
        IF(ISIGN) 410,390,390
390     WI=-WI
410   W2R=WR*WR-WI*WI
      W2I=2.*WR*WI
      W3R=W2R*WR-W2I*WI
      W3I=W2R*WI+W2I*WR
      KMIN=1+IPAR*M
        GO TO 440
420   KMIN=1
440   KDIF=IPAR*MMAX
450   KSTEP=4*KDIF
      IF(KSTEP-64)460,460,530
```

```
460   DO 520 K1=KMIN,192,KSTEP
      K2=K1+KDIF
      K3=K2+KDIF
      K4=K3+KDIF
      IF(MMAX-2)470,470,480
470   U1R=DATA(K1)/2+DATA(K2)/2
      U1I=DATA(K1+1)/2+DATA(K2+1)/2
      U2R=DATA(K3)/2+DATA(K4)/2
      U2I=DATA(K3+1)/2+DATA(K4+1)/2
      U3R=DATA(K1)/2-DATA(K2)/2
      U3I=DATA(K1+1)/2-DATA(K2+1)/2
      IF(ISIGN)471,472,472
471   U4R=DATA(K3+1)/2-DATA(K4+1)/2
      U4I=DATA(K4)/2-DATA(K3)/2
      GO TO 510
472   U4R=DATA(K4+1)/2-DATA(K3+1)/2
      U4I=DATA(K3)/2-DATA(K4)/2
      GO TO 510
480   T2R=W2R*DATA(K2)/2-W2I*DATA(K2+1)/2
      T2I=W2R*DATA(K2+1)/2+W2I*DATA(K2)/2
      T3R=WR*DATA(K3)/2-WI*DATA(K3+1)/2
      T3I=WR*DATA(K3+1)/2+WI*DATA(K3)/2
      T4R=W3R*DATA(K4)/2-W3I*DATA(K4+1)/2
      T4I=W3R*DATA(K4+1)/2+W3I*DATA(K4)/2
      U1R=DATA(K1)/2+T2R
      U1I=DATA(K1+1)/2+T2I
      U2R=T3R+T4R
      U2I=T3I+T4I
      U3R=DATA(K1)/2-T2R
      U3I=DATA(K1+1)/2-T2I
      IF(ISIGN)490,500,500
490   U4R=T3I-T4I
      U4I=T4R-T3R
      GO TO 510
500   U4R=T4I-T3I
      U4I=T3R-T4R
510   DATA(K1)=U1R+U2R
      DATA(K1+1)=U1I+U2I
      DATA(K2)=U3R+U4R
      DATA(K2+1)=U3I+U4I
      DATA(K3)=U1R-U2R
      DATA(K3+1)=U1I-U2I
      DATA(K4)=U3R-U4R
520   DATA(K4+1)=U3I-U4I
      KDIF=KSTEP
        KMIN=4*KMIN-3
      GO TO 450
```

```
530     M=M+LMAX
        IF(M-MMAX)540,540,570
540     IF(ISIGN)550,560,560
550     TEMPR=WR
        WR=(WR+WI)*RTHLF
        WI=(WI-TEMPR)*RTHLF
        GO TO 410
560     TEMPR=WR
        WR=(WR-WI)*RTHLF
        WI=(TEMPR+WI)*RTHLF
        GO TO 410
570     CONTINUE
        IPAR=3-IPAR
        MMAX=MMAX+MMAX
        GO TO 360
C
C       MAIN LOOP FOR FACTORS NOT EQUAL TO TWO.
C       W=EXP(ISIGN*2*PI*SQRT(-1)*(J1+J2-I3-1)/IFP2)
C
600       WSTPI=-.8660254
          IF(ISIGN)612,611,611
611       WSTPI=.8660254
612       L123=0
        DO 650 J1=1,64,2
          L123=L123+1
          WR=BCOS(L123)
          WI=BSIN(L123)
          IF(ISIGN) 614,613,613
613       WI=-WI
614       SR=WR*DATA(J1+128)+DATA(J1+64)/2
        SI=WR*DATA(J1+129)+DATA(J1+65)/2
          A1=-DATA(J1+128)/2+DATA(J1)/2
          A2=-DATA(J1+129)/2+DATA(J1+1)/2
        WORK(1)=WR*SR-WI*SI+A1
        WORK(2)=WI*SR+WR*SI+A2
          WTEMP=WR*WSTPI
          WR=-.5*WR-WI*WSTPI
          WI=-.5*WI+WTEMP
          SR=WR*DATA(J1+128)+DATA(J1+64)/2
          SI=WR*DATA(J1+129)+DATA(J1+65)/2
          WORK(3)=WR*SR-WI*SI+A1
          WORK(4)=WI*SR+WR*SI+A2
          WTEMP=WR*WSTPI
          WR=-.5*WR-WI*WSTPI
          WI=-.5*WI+WTEMP
          SR=WR*DATA(J1+128)+DATA(J1+64)/2
          SI=WR*DATA(J1+129)+DATA(J1+65)/2
          DATA(J1+128)=WR*SR-WI*SI+A1
          DATA(J1+129)=WI*SR+WR*SI+A2
          DATA(J1)=WORK(1)
          DATA(J1+1)=WORK(2)
          DATA(J1+64)=WORK(3)
650       DATA(J1+65)=WORK(4)
          RETURN
          END
```

END

FILMED